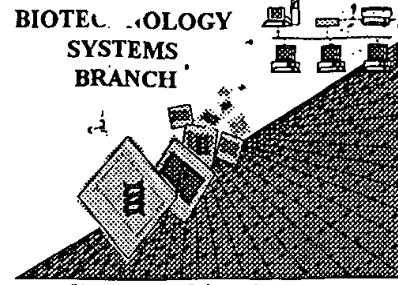


RAW SEQUENCE LISTING ERROR REPORT



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/806,232A
Source: Pat/09
Date Processed by STIC: 8/2/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be downloaded from the USPTO website at the following address:
<http://www.uspto.gov/web/offices/pac/checker>

BEST AVAILABLE COPY

PCT09

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/806,232A

DATE: 08/02/2001
TIME: 16:38:05

Input Set : A:\1241.18 Sequence.txt
Output Set: N:\CRF3\08022001\I806232A.raw

3 <110> APPLICANT: Seiki Motoharu
5 <120> TITLE OF INVENTION: DNA CODING FOR NOVEL POLYPEPTIDE
7 <130> FILE REFERENCE: 1241.18
9 <140> CURRENT APPLICATION NUMBER: US 09/806,232A
C--> 10 <141> CURRENT FILING DATE: 2001-07-18 *7/15/07*
W--> 12 <140> CURRENT APPLICATION NUMBER: PCT/JP99/05349A *7/15/97*
C--> 13 <141> CURRENT FILING DATE: 1999-09-29
15 <150> PRIOR APPLICATION NUMBER: JP10-276258
16 <151> PRIOR FILING DATE: 1998-09-29
18 <150> PRIOR APPLICATION NUMBER: JP10-291505
19 <151> PRIOR FILING DATE: 1998-09-29
21 <160> NUMBER OF SEQ ID NOS: 22
23 <170> SOFTWARE: PatentIn Ver. 2.0

Does Not Comply
Corrected Diskette Needed

These are
prior data

ERRORED SEQUENCES

481 <210> SEQ ID NO: 4
482 <211> LENGTH: 2423 2438 *what are these? 2438 shown (p. 4)*
483 <212> TYPE: DNA
484 <213> ORGANISM: Homo sapiens
486 <220> FEATURE:
487 <221> NAME/KEY: CDS
488 <222> LOCATION: (100)..(1917)
490 <400> SEQUENCE: 4
491 cccggggggg cggccggggag agcggaggggc gcccggctgc ggaacgcgaa gcggaggggcg 60
493 cgggaccctg cacgcccccc gccccccat gtgagcgcc atg cgg cgc cgc gca 114
494 Met Arg Arg Arg Ala
495 1 *5 make up under*
E--> 496 5
498 gcc cgg gga ccc ggc ccg ccc cca ggg ccc gga ctc tcg cgg ctg 162
499 Ala Arg Gly Pro Gly Pro Pro Pro Gly Pro Gly Leu Ser Arg Leu
W--> 500 10 15 20
502 ccg ctg ctg ccg ctg ctg ctg ctg ctg ctg ggg acc cgc 210
503 Pro Leu Leu Pro Leu Leu Leu Leu Ala Leu Gly Thr Arg
W--> 504 25 30 35
506 ggg ggc tgc gcc gcg ccg gaa ccc gcg cgg cgc gcc gag gac ctc agc 258
507 Gly Gly Cys Ala Ala Pro Glu Pro Ala Arg Arg Ala Glu Asp Leu Ser
W--> 508 40 45 50
510 ctg gga gtg gag tgg cta agc agg ttc ggt tac ctg ccc ccg gct gac 306
511 Leu Gly Val Glu Trp Leu Ser Arg Phe Gly Tyr Leu Pro Pro Ala Asp
W--> 512 55 60 65
514 ccc aca aca ggg cag ctg cag acg caa gag gag ctg tct aag gcc atc 354
515 Pro Thr Thr Gly Gln Leu Gln Thr Gln Glu Glu Leu Ser Lys Ala Ile
W--> 516 70 75 80 85
518 aca gcc atg cag cag ttt ggt ggc ctg gag gcc acc ggc atc ctg gac 402
519 Thr Ala Met Gln Gln Phe Gly Gly Leu Glu Ala Thr Gly Ile Leu Asp

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/806,232A

DATE: 08/02/2001
TIME: 16:38:05

Input Set : A:\1241.18 Sequence.txt
Output Set: N:\CRF3\08022001\I806232A.raw

W--> 520 90 95 100
 522 gag gcc acc ctg gcc ctg atg aaa acc cca cgc tgc tcc ctg cca gac 450
 523 Glu Ala Thr Leu Ala Leu Met Lys Thr Pro Arg Cys Ser Leu Pro Asp
 W--> 524 105 110 115
 526 ctc cct gtc ctg acc cag gct cgc agg aga cgc cag gct cca gcc ccc 498
 527 Leu Pro Val Leu Thr Gln Ala Arg Arg Arg Arg Gln Ala Pro Ala Pro
 W--> 528 120 125 130
 530 acc aag tgg aac aag agg aac ctg tcg tgg agg gtc cgg acg ttc cca 546
 531 Thr Lys Trp Asn Lys Arg Asn Leu Ser Trp Arg Val Arg Thr Phe Pro
 W--> 532 135 140 145
 534 cgg gac tca cca ctg ggg cac gac acg gtg cgt gca ctc atg tac tac 594
 535 Arg Asp Ser Pro Leu Gly His Asp Thr Val Arg Ala Leu Met Tyr Tyr
 W--> 536 150 155 160 165
 538 gcc ctc aag gtc tgg agc gac att gcg ccc ctg aac ttc cac gag gtg 642
 539 Ala Leu Lys Val Trp Ser Asp Ile Ala Pro Leu Asn Phe His Glu Val
 W--> 540 170 175 180
 542 ggc ggc agc acc gcc gac atc cag atc gac ttc tcc aag gcc gac cat 690
 543 Ala Gly Ser Thr Ala Asp Ile Gln Ile Asp Phe Ser Lys Ala Asp His
 W--> 544 185 190 195
 546 aac gac ggc tac ccc ttc gac ggc ccc ggc acc gtg gcc cac gcc 738
 547 Asn Asp Gly Tyr Pro Phe Asp Gly Pro Gly Gly Thr Val Ala His Ala
 W--> 548 200 205 210
 550 ttc ttc ccc ggc cac cac cac acc gcc ggg gac acc cac ttt gac gat 786
 551 Phe Phe Pro Gly His His Thr Ala Gly Asp Thr His Phe Asp Asp
 W--> 552 215 220 225
 554 gac gag gcc tgg acc ttc cgc tcc tcg gat gcc cac ggg atg gac ctg 834
 555 Asp Glu Ala Trp Thr Phe Arg Ser Ser Asp Ala His Gly Met Asp Leu
 W--> 556 230 235 240 245
 558 ttt gca gtg gtc cac gag ttt ggc cac att ggg tta agc cat 882
 559 Phe Ala Val Ala Val His Glu Phe Gly His Ala Ile Gly Leu Ser His
 W--> 560 250 255 260
 562 gtg gcc gct gca cac tcc atc atg cgg cgg tac tac cag ggc ccc gtg 930
 563 Val Ala Ala Ala His Ser Ile Met Arg Pro Tyr Tyr Gln Gly Pro Val
 W--> 564 265 270 275
 566 ggt gac ccg ctg cgc tac ggg ctc ccc tac gag gac aag gtg cgc gtc 978
 567 Gly Asp Pro Leu Arg Tyr Gly Leu Pro Tyr Glu Asp Lys Val Arg Val
 W--> 568 280 285 290
 570 tgg cag ctg tac ggt gtg cgg gag tct gtg tct ccc acg ggc cag ccc 1026
 571 Trp Gln Leu Tyr Gly Val Arg Glu Ser Val Ser Pro Thr Ala Gln Pro
 W--> 572 295 300 305
 574 gag gag cct ccc ctg cgg gag ccc cca gac aac cgg tcc agc gcc 1074
 575 Glu Glu Pro Pro Leu Leu Pro Glu Pro Pro Asp Asn Arg Ser Ser Ala
 W--> 576 310 315 320 325
 578 ccg ccc agg aag gac gtg ccc cac aga tgc agc act cac ttt gac gcg 1122
 579 Pro Pro Arg Lys Asp Val Pro His Arg Cys Ser Thr His Phe Asp Ala
 W--> 580 330 335 340
 582 gtg gcc cag atc cgg ggt gaa gct ttc ttc ttc aaa ggc aag tac ttc 1170
 583 Val Ala Gln Ile Arg Gly Glu Ala Phe Phe Phe Lys Gly Lys Tyr Phe
 W--> 584 345 350 355

OK

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/806,232A

DATE: 08/02/2001
TIME: 16:38:05

Input Set : A:\1241.18 Sequence.txt
Output Set: N:\CRF3\08022001\I806232A.raw

586 tgg cgg ctg acg cgg gac cgg cac ctg gtg tcc ctg cag cgg gca cag 1218
587 Trp Arg Leu Thr Arg Asp Arg His Leu Val Ser Leu Gln Pro Ala Gln
W--> 588 360 365 370
590 atg cac cgc ttc tgg cgg ggc ctg cgg cac ctg gac agc gtg gac 1266
591 Met His Arg Phe Trp Arg Gly Leu Pro Leu His Leu Asp Ser Val Asp
W--> 592 375 380 385
594 gcc gtg tac gag cgc acc agc gac cac aag atc gtc ttc ttt aaa gga 1314
595 Ala Val Tyr Glu Arg Thr Ser Asp His Lys Ile Val Phe Phe Lys Gly
W--> 596 390 395 400 405
598 gac agg tac tgg gtg ttc aag gac aat aac gta gag gaa gga tac cgg 1362
599 Asp Arg Tyr Trp Val Phe Lys Asp Asn Asn Val Glu Glu Gly Tyr Pro
W--> 600 410 415 420
602 cgc ccc gtc tcc gac ttc agc ctc cgg cct ggc ggc atc gac gct gcc 1410
603 Arg Pro Val Ser Asp Phe Ser Leu Pro Pro Gly Gly Ile Asp Ala Ala
W--> 604 425 430 435
606 ttc tcc tgg gcc cac aat gac agg act tat ttc ttt aag gac cag ctg 1458
607 Phe Ser Trp Ala His Asn Asp Arg Thr Tyr Phe Phe Lys Asp Gln Leu
W--> 608 440 445 450
610 tac tgg cgc tac gat gac cac acg agg cac atg gac ccc ggc tac ccc 1506
611 Tyr Trp Arg Tyr Asp Asp His Thr Arg His Met Asp Pro Gly Tyr Pro
W--> 612 455 460 465
614 gcc cag agc ccc ctg tgg agg ggt gtc ccc agc acg ctg gac gac gcc 1554
615 Ala Gln Ser Pro Leu Trp Arg Gly Val Pro Ser Thr Leu Asp Asp Ala
W--> 616 470 475 480 485
618 atg cgc tgg tcc gac ggt gcc tcc tac ttc ttc cgt ggc cag gag tac 1602
619 Met Arg Trp Ser Asp Gly Ala Ser Tyr Phe Phe Arg Gly Gln Glu Tyr
W--> 620 490 495 500
622 tgg aaa gtg ctg gat ggc gag ctg gag gtt gca ccc ggg tac cca cag 1650
623 Trp Lys Val Leu Asp Gly Glu Leu Glu Val Ala Pro Gly Tyr Pro Gln
W--> 624 505 510 515
626 tcc acg gcc cgg gac tgg ctg gtt ggt gca gac tca cag gcc gat gga 1698
627 Ser Thr Ala Arg Asp Trp Leu Val Cys Gly Asp Ser Gln Ala Asp Gly
W--> 628 520 525 530
630 tct gtg gct gcg ggc gtg gac gcg gca gag ggg ccc cgc gcc cct cca 1746
631 Ser Val Ala Ala Gly Val Asp Ala Ala Glu Gly Pro Arg Ala Pro Pro
W--> 632 535 540 545
634 gga caa cat gac cag agc cgc tcg gag gac ggt tac gag gtc tgc tca 1794
635 Gly Gln His Asp Gln Ser Arg Ser Glu Asp Gly Tyr Glu Val Cys Ser
W--> 636 550 555 560 565
638 tgc acc tct ggg gca tcc tct ccc cgg ggg gcc cca ggc cca ctg gtg 1842
639 Cys Thr Ser Gly Ala Ser Ser Pro Pro Gly Ala Pro Gly Pro Leu Val
W--> 640 570 575 580
642 gct gcc acc atg ctg ctg ctg cgg cca ctg tca cca ggc gcc ctg 1890
643 Ala Ala Thr Met Leu Leu Leu Pro Pro Leu Ser Pro Gly Ala Leu
W--> 644 585 590 595
646 tgg aca gcg gcc cag gcc ctg acg cta tgacacacag cgcgagccca 1937
647 Trp Thr Ala Ala Gln Ala Leu Thr Leu
W--> 648 600 605
650 tgagaggaca gaggcggtgg gacagcctgg ccacagaggg caaggactgt gcccggagtcc 1997

OK

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/806,232A

DATE: 08/02/2001
TIME: 16:38:05

Input Set : A:\1241.18 Sequence.txt
Output Set: N:\CRF3\08022001\I806232A.raw

652 ctgggggagg tgctggcgcg ggtatgaggac gggccaccct ggcaccggaa ggcagcaga 2057
654 gggcacggcc cgccagggtc gggcagggtc aggtggcaag gacggagctg tcccttagtg 2117
656 agggactgtg ttgactgacg agccgagggg tggccgctcc agaagggtgc ccagtcaaggc 2177
658 cgcacccgccc ccagccctcc ctggccctgg agggagcatc tcgggctggg ggcccaaaaa 2237
660 tctctgtgcc ggccacca accccaccca cactgctgcc tggtgctccc gccggccac 2297
662 agggcctccg tccccaggtc cccagtggg cagccctccc cacagacgag ccccccacat 2357
664 ggtgccgggg cacgtcccccc ctgtgacgctg ttccagacca acatgaccc tccctgcttt 2417
E--> 666 gtaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa a 2438 ↙

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/806,232A

DATE: 08/02/2001
TIME: 16:38:06

Input Set : A:\1241.18 Sequence.txt
Output Set: N:\CRF3\08022001\I806232A.raw

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:12 M:280 W: Numeric Identifier already exists, <140> found multiple times
L:12 M:281 W: Numeric Fields not Ordered, <140> not ordered!.
L:12 M:270 C: Current Application Number differs, Replaced Current Application Number
L:13 M:280 W: Numeric Identifier already exists, <141> found multiple times
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:496 M:254 E: No. of Bases conflict, LENGTH:Input:5 Counted:114 SEQ:4
L:500 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:504 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:508 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:512 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:516 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:520 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:524 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:528 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:532 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:536 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:540 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:544 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:548 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:552 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:556 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:560 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:564 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:568 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:572 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:576 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:580 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:584 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:588 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:592 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:596 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:600 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:604 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:608 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:612 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:616 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:620 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:624 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:628 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:632 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:636 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:640 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:644 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:648 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:4
L:666 M:252 E: No. of Seq. differs, <211>LENGTH:Input:2423 Found:2438 SEQ:4